

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 17, 2003, 08:49:19 ; Search time 159 Seconds  
(without alignments)  
3154.732 Million cell updates/sec

Title: US-10-010-227-3  
Perfect score: 4055  
Sequence: 1 MPAESTPQTLXDKVLQAHV.....KAVPVTNRGEKEKPLEW 778

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

## Database :

Pending Patents\_AA Main:\*

1: /cgn2\_6/ptodata/2/paa/US06 COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06 COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07 COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08 COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081 COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082 COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083 COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084 COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085 COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086 COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087 COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088 COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089 COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090 COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091 COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092 COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093 COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094 COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095 COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096 COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097 COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US098 COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US099 COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US100 COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US101 COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US102 COMB.pep.\*  
27: /cgn2\_6/ptodata/2/paa/US60 COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4055	100.0	778	24	US-10-010-227-3
2	3397	83.8	840	27	US-60-360-039-1233
3	2456.5	60.6	779	27	US-60-360-039-11952
4	2451.5	60.5	794	25	US-10-179-131-8614
5	2076	51.2	644	21	US-09-791-537-121852
6	2068.5	51.0	875	27	US-60-360-039-7864

7	2055	50.7	695	27	US-60-360-039-9965	Sequence 9965, Ap
8	2018.5	49.0	710	27	US-60-360-039-15849	Sequence 15849, A
9	1995	48.2	752	27	US-60-360-039-15479	Sequence 15479, A
10	1978	48.8	755	27	US-60-360-039-9413	Sequence 9413, Ap
11	1978	48.4	749	27	US-60-360-039-17883	Sequence 17883, A
12	1953.5	48.4	681	27	US-60-360-039-9308	Sequence 9308, Ap
13	1955.5	48.2	749	27	US-60-360-039-8616	Sequence 8616, Ap
14	1891	46.6	740	27	US-60-360-039-7437	Sequence 7437, Ap
15	1881	46.4	729	27	US-60-360-039-4678	Sequence 4678, Ap
16	1861	45.9	686	27	US-60-360-039-7055	Sequence 7055, Ap
17	1857	45.8	672	27	US-60-360-039-9059	Sequence 9059, Ap
18	1711.5	42.2	678	27	US-60-360-039-9059	Sequence 9059, Ap
19	1637.5	41.9	676	27	US-60-360-039-10463	Sequence 10463, A
20	1682.5	41.2	646	27	US-60-360-039-16643	Sequence 16643, A
21	1670.5	41.5	679	27	US-60-360-039-16436	Sequence 16436, A
22	1666.5	41.1	644	27	PCT-US02-03987-5730	Sequence 14041, A
23	1666.5	41.1	644	22	US-09-815-242-5730	Sequence 5730, Ap
24	1666.5	41.1	644	24	US-10-072-851-5730	Sequence 5730, Ap
25	1666.5	41.1	644	27	US-60-242-578-1009	Sequence 1009, Ap
26	1666.5	41.1	644	27	US-60-253-625-2353	Sequence 2353, Ap
27	1666.5	41.1	644	27	US-60-257-931-3312	Sequence 3312, Ap
28	1666.5	41.1	644	27	US-60-269-308-4336	Sequence 4336, Ap
29	1655.5	40.8	517	18	US-09-417-507-37934	Sequence 37934, A
30	1595.5	39.3	668	27	US-60-360-039-8819	Sequence 8819, Ap
31	1586	39.1	469	27	US-60-360-039-15002	Sequence 15002, A
32	1555	38.3	534	24	US-10-015-127-11081	Sequence 10081, A
33	1548.5	38.2	469	21	US-09-791-537-20646	Sequence 20646, A
34	1545.5	38.1	469	21	US-09-791-537-33097	Sequence 33097, A
35	1529.5	37.7	471	27	US-60-360-039-19706	Sequence 19706, A
36	1520.5	37.5	526	17	US-09-328-352-8024	Sequence 8024, Ap
37	1508	37.2	672	27	US-60-360-039-7323	Sequence 7323, Ap
38	1507.5	37.2	474	1	PCT-US02-03987-5127	Sequence 5127, Ap
39	1507.5	37.2	474	22	US-09-815-242-5127	Sequence 5127, Ap
40	1507.5	37.2	474	24	US-10-072-851-5127	Sequence 5127, Ap
41	1507.5	37.2	474	27	US-60-253-625-2683	Sequence 2683, Ap
42	1507.5	37.2	474	27	US-60-257-931-3511	Sequence 3511, Ap
43	1507.5	37.2	474	27	US-60-269-308-4533	Sequence 4533, Ap
44	1507.5	37.2	474	27	US-60-269-308-4533	Sequence 4533, Ap
45	1507.5	37.2	480	16	US-09-252-9914-28668	Sequence 28668, A
46	1494	36.8	670	27	US-60-360-039-94565	Sequence 94565, Ap
47	1483.5	36.6	466	21	US-09-791-537-15038	Sequence 15038, A
48	1483.5	36.6	481	18	US-09-489-039A-12322	Sequence 12322, A
49	1479.5	36.5	466	21	US-09-791-537-26234	Sequence 26234, A
50	1479.5	36.5	466	21	US-09-791-537-84958	Sequence 84958, A
51	1479.5	36.5	466	21	US-09-791-537-84958	Sequence 84958, A
52	1479.5	36.5	466	21	US-09-791-537-84958	Sequence 84958, A
53	1471	36.3	469	1	PCT-US02-03987-11164	Sequence 11164, A
54	1471	36.3	469	21	US-09-815-242-11164	Sequence 11164, A
55	1471	36.3	469	21	US-09-815-242-11164	Sequence 11164, A
56	1471	36.3	469	21	US-09-815-242-11164	Sequence 11164, A
57	1471	36.3	469	21	US-09-815-242-11164	Sequence 11164, A
58	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
59	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
60	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
61	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
62	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
63	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
64	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
65	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
66	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
67	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
68	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
69	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
70	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
71	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
72	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
73	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
74	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
75	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
76	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
77	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
78	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
79	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
80	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
81	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
82	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
83	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
84	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
85	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
86	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
87	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
88	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
89	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
90	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
91	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
92	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
93	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
94	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
95	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
96	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
97	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
98	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
99	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap
100	1467.5	36.2	470	16	US-09-252-691-9422	Sequence 9422, Ap

80 1311.5 32.3 470 21 US-09-791-537-58954 Sequence 55954, A  
81 1304.5 32.2 464 21 US-09-791-537-108096 Sequence 108096, A  
82 1300.5 32.1 461 18 US-09-450-969-6326 Sequence 6326, Ap  
83 1290.5 31.8 466 21 US-09-791-537-90886 Sequence 90886, A  
84 1282 31.6 456 1 PCT-US02-03987-12156 Sequence 12156, A  
85 1282 31.6 456 1 PCT-US02-03987-12928 Sequence 12928, A  
86 1282 31.6 456 12 US-08-827-356-5398 Sequence 5398, Ap  
87 1282 31.6 456 20 US-09-611-523-6093 Sequence 6093, Ap  
88 1282 31.6 456 22 US-09-815-242-12156 Sequence 12156, A  
89 1282 31.6 456 22 US-09-815-242-12928 Sequence 12928, A  
90 1282 31.6 456 24 US-10-072-851-12156 Sequence 12156, A  
91 1282 31.6 456 24 US-10-072-851-12928 Sequence 12928, A  
92 1244 30.7 474 19 US-09-540-2098-10069 Sequence 10069, A  
93 1209.5 29.8 472 21 US-09-791-537-4868 Sequence 4868, Ap  
94 1181.5 29.1 468 21 US-09-791-537-15039 Sequence 15039, A  
95 1178.5 29.1 467 27 US-60-360-039-19849 Sequence 19849, A  
96 1172.5 28.9 463 27 US-60-360-039-18971 Sequence 18971, A  
97 1141.5 28.2 460 21 US-09-791-537-81237 Sequence 81237, A  
98 1132.5 27.9 460 27 US-60-360-039-18414 Sequence 18414, A  
99 1114 27.5 459 21 US-09-791-537-131423 Sequence 131423, A  
100 1076 26.5 287 18 US-09-417-507-37932 Sequence 37932, A

ALIGNMENTS

RESULT 1  
US-10-010-227-3  
; Sequence 3, Application US/10010227  
; GENERAL INFORMATION:  
; APPLICANT: Hamer, Lisbeth  
; APPLICANT: Adachi, Kiichi  
; APPLICANT: DeZwaan, Todd M  
; APPLICANT: Lo, Sze Chung C  
; APPLICANT: Montenegro-Chamorro, Maria V  
; APPLICANT: Frank, Sheryl A  
; APPLICANT: Darveaux, Blaise A  
; APPLICANT: Mahanty, Sanjoy K  
; APPLICANT: Heiniger, Ryan W  
; APPLICANT: Skalchunes, Amy R  
; APPLICANT: Pan, Huaqin  
; APPLICANT: Tarpey, Rex  
; APPLICANT: Shuster, Jeffrey R  
; APPLICANT: Tanzer, Matthew M  
; TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF INHIBITORS OF 3-ISOPROPYLMALATE  
; TITLE OF INVENTION: DEHYDRATASE AS ANTIBIOTICS  
; FILE REFERENCE: 2131US  
; CURRENT APPLICATION NUMBER: US/10/010,227  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 778  
; TYPE: PRT  
; ORGANISM: Magnaporthe grisea  
US-10-010-227-3

Query Match 100.0%; Score 4055; DB 24; Length 778;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 778; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MFGAESTPOTLYDKVLOAHVDEKLDGTLLYIDRHLVHEVTSPOAFEGRLNAGRKVRP 60  
Db 1 MFGAESTPOTLYDKVLOAHVDEKLDGTLLYIDRHLVHEVTSPOAFEGRLNAGRKVRP 60  
Qy 61 DCTLATTDHNVPTTSRKALKDIASFKEKEDDSRTQCVTLEENKVEFGVTYFGLSDKRGQIV 120  
Db 61 DCTLATTDHNVPTTSRKALKDIASFKEKEDDSRTQCVTLEENKVEFGVTYFGLSDKRGQIV 120  
Qy 121 HVIGPEQGTTLGTTVCGDSHTSTHGAFGALAFGICTSEVHEVLATQCLITKRSKNMRI 180  
Db 121 HVIGPEQGTTLGTTVCGDSHTSTHGAFGALAFGICTSEVHEVLATQCLITKRSKNMRI 180

Qy 181 QVDGELAPGVSSKDVVLHAIGIITAGTGAVIEFCGVSIRLSMEARMSICNMSIEGGA 240  
Db 181 QVDGELAPGVSSKDVVLHAIGIITAGTGAVIEFCGVSIRLSMEARMSICNMSIEGGA 240  
Qy 241 RAGWVAPDEITFEYLKGRPLAPKYDSPWHKATQWKNLQSDPGAKYDIDVFTDAKDIVP 300  
Db 241 RAGWVAPDEITFEYLKGRPLAPKYDSPWHKATQWKNLQSDPGAKYDIDVFTDAKDIVP 300  
Qy 301 TLTWGTSPEDVVPITGVVPDPETPATEAKADGRMLQYMGKAGTPMEDIPVDKVFIS 360  
Db 301 TLTWGTSPEDVVPITGVVPDPETPATEAKADGRMLQYMGKAGTPMEDIPVDKVFIS 360  
Qy 361 CTNSRIEDLRAAAAIVVKRKKAPNVKSAWVPGSLVKTQAEELGDKIFEAGFEWREA 420  
Db 361 CTNSRIEDLRAAAAIVVKRKKAPNVKSAWVPGSLVKTQAEELGDKIFEAGFEWREA 420  
Qy 421 GCSMCLGNPNPILAPOERCASSTNRNFRGCGAGRTHLMSVMAAAAGIVGKLADVRKL 480  
Db 421 GCSMCLGNPNPILAPOERCASSTNRNFRGCGAGRTHLMSVMAAAAGIVGKLADVRKL 480  
Qy 481 TDYKASPHIAAYQKSTVTKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSA 540  
Db 481 TDYKASPHIAAYQKSTVTKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSA 540  
Qy 541 PKFTILKGIAAPLEKANVDTDAIIPKQFLTKITKRTGLGNALFYEMRPNEDGTEKSDFVLN 600  
Db 541 PKFTILKGIAAPLEKANVDTDAIIPKQFLTKITKRTGLGNALFYEMRPNEDGTEKSDFVLN 600  
Qy 601 KEPYRKASILVCTGANFCGSSRREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGMLPIP 660  
Db 601 KEPYRKASILVCTGANFCGSSRREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGMLPIP 660  
Qy 661 IKDQAIEAIAAEARAGKEIEVDLPNQLIKNATGETICTFEVEEPRKHCLVNGLDDIGLT 720  
Db 661 IKDQAIEAIAAEARAGKEIEVDLPNQLIKNATGETICTFEVEEPRKHCLVNGLDDIGLT 720  
Qy 721 MOMEDKTAEPFAKMTRETPWLDGTGYLKRGGGKLAAKAVPVTNNRGEKKEPLEW 778  
Db 721 MOMEDKTAEPFAKMTRETPWLDGTGYLKRGGGKLAAKAVPVTNNRGEKKEPLEW 778

RESULT 2  
US-60-360-039-3233  
; Sequence 3233, Application US/60360039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 3233  
; LENGTH: 840  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(840)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-3233

Query Match 83.8%; Score 3397; DB 27; Length 840;  
Best Local Similarity 77.8%; Pred. No. 0;  
Matches 655; Conservative 53; Mismatches 60; Indels 74; Gaps 5;  
Qy 9 QTLYDKVLOAHVDEKLDGTLLYIDRHLVHEVTS----- 44  
Db 1 RTLYDKVFOAHIVDEKLDGTLLYIDRHLVHEVTSVRSLNPHAGELFRSPVPSTDPKS 60

```

QY      45 -OAEGLRNAGKVRRPDCTLATTDH----- 69
      61 IQAEGLRNAGKVRRPDCTLATTDHVSMSLPLFEYFMAAHAFAFRXYLAPSSLCG 120
QY      70 -----NVPTTSRKALKDIAFIKEDSRTOCVTLLENVKEFGVTYFGLSDRKGI 119
      121 TASXPRTRCRWVPTTSRKALKDIAFIKEDSRTOCVTLLENVKKFGITTYFGLSDRKGI 180
QY      120 VHVIVGPGGFTLPGTTVVCGBSHTSTHGAFGALAFGIISTSEVENHLATQCLITRSKMR 179
      181 VHVIVGPGGFTLPGTTVVCGBSHTSTHGAFGALAFGIISTSEVENHLATQCLITRSKMR 240
QY      180 IOVDGELAPGVSSKDVVLAHAIIGTAGTGAVIEFCGSVIRSLSEARMSICNMSIEGG 239
      241 VOVDGELAPGVSSKDVVLAHAIIGTAGTGAVIEFCGSVIRSLSEARMSICNMSIAG 300
QY      240 ARAGVAPDEITFEYLKGRPLAPKXDSPEMHKATQYWKLOSDPAKXIDIVFIDAKIV 299
      301 ARAGVAPDEITFEYLKGRPLAPRGSEENKAVVYWKSLASDPKXIDIVFIDAKI 360
QY      300 PTLWNGSPEDVVPITGVVPDETFATKADGRMLQYMGKAGTPEMEDIPTDVKVF 359
      361 PTLWNGSPEDVVPITGVVPDETFATKADGRMLQYMGKAGTPEMEDIPTDVKVF 420
QY      360 SCTNSRIEDLRAAAAVVGRKAPNVKSGAMVPGSLVKTQAEEDGLDKIEBAGFEWRE 419
      421 SCTNSRIEDLRAAAAVVGRKAPNVKSGAMVPGSLVKTQAEEDGLDKIEBAGFEWRE 480
QY      420 AGCSMCLGMPDILAPORCASTSNRNFEGROGAGGRTHLMSPVMAAAAGIVGLADVRK 479
      481 AGCSMCLGMPDILAPORCASTSNRNFEGROGAGGRTHLMSPVMAAAAGIVGLADVRK 540
QY      480 LTDVKSAPHIAV---OKSTVTKPHVDERINODAEHKOIADIPEDNNGPHNTSASVCT 536
      541 LTDVKSAPHIAV---OKSTVTKPHVDERINODAEHKOIADIPEDNNGPHNTSASVCT 598
QY      537 SAGLPKFTILKGIAPLEKANVDTDAIIPKQFLTKIKRTGLGNALFYEMRFNEDGTEKSD 596
      599 SAGLPKFTILKGIAPLEKANVDTDAIIPKQFLTKIKRTGLGNALFYEMRFNEDGTEKSD 658
QY      597 FVLNKEPRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIASPADIFFNNSFKMG 656
      659 FVLNKEPRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIASPADIFFNNSFKMG 718
QY      657 LPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRKICLVNGLDD 716
      719 LPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRKICLVNGLDD 778
QY      717 IGLTWQMEDKIAFEAKMTRETPMLDGTGYLKRKGCGGLAKAAVAVPTTNRGEEKKEPL 776
      779 IGLTWQMEDKIAFEAKMTRETPMLDGTGYLKRKGCGGLAKAAVAVPTTNRGEEKKEPL 838
QY      777 EW 778
      839 EW 840

```

RESULT 3  
 US-60-360-039-21952  
 ; Sequence 21952, Application US/60360039  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Chen, Xianfeng  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; FILE REFERENCE: 38-10(52052)A  
 ; CURRENT APPLICATION NUMBER: US/60/360, 039  
 ; CURRENT FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374

```

; SEQ ID NO 21952
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-60-360-039-21952

Query Match      60.6%; Score 2456.5; DB 27; Length 779;
Best Local Similarity 63.1%; Pred. No. 5,3e-226;
Matches 486; Conservative 93; Mismatches 152; Indels 39; Gaps 10;

QY      8 POTLYDKVLAHVDEKDGTVLLYIDRLVHEVTSPOAFEGELRNAGKVRRPDCTLAT 67
      9 PRTLYDKVLAHVDEKDGTVLLYIDRLVHEVTSPOAFEGELRNAGKVRRPDCTLAT 68
QY      68 DHNVPTTSRKALKDIAFIKEDSRTOCVTLLENVKEFGVTYFGLSDRKGIIVHIGPEQ 127
      69 DHNVPTTSRKALKDIAFIKEDSRTOCVTLLENVKEFGVTYFGLSDRKGIIVHIGPEQ 128
QY      128 GFTLPGTTVVCGBSHTSTHGAFGALAFGIISTSEVENHLATQCLITRSKMRIOVDGELA 187
      129 GFTLPGTTVVCGBSHTSTHGAFGALAFGIISTSEVENHLATQCLITRSKMRIOVDGELA 188
QY      188 PGVSKDVVLAHAIIGTAGTGAVIEFCGSVIRSLSEARMSICNMSIEGARAGVAP 247
      189 PGVSKDVVLAHAIIGTAGTGAVIEFCGSVIRSLSEARMSICNMSIEGARAGVAP 248
QY      248 DEITFEYLKGRPLAPKXDSPEMHKATQYWKLOSDPAKXIDIVFIDAKIVPILTWGTS 307
      249 DEITFEYLKGRPLAPKXDSPEMHKATQYWKLOSDPAKXIDIVFIDAKIVPILTWGTS 306
QY      308 PEDVVPITGVVPDETFATKADGRMLQYMGKAGTPEMEDIPTDVKVFISGCTNSRIE 367
      309 PEDVVPITGVVPDETFATKADGRMLQYMGKAGTPEMEDIPTDVKVFISGCTNSRIE 366
QY      307 PODALPTIGSVDPKXVTDPIKSGNERALAMGLEPNTPLSLIVDVKVFISGCTNSRIE 366
      308 PODALPTIGSVDPKXVTDPIKSGNERALAMGLEPNTPLSLIVDVKVFISGCTNSRIE 367
QY      368 DLRAAAAYVGRKAPNVKSGAMVPGSLVKTQAEEDGLDKIEBAGFEWREAGCSMCLG 427
      369 DLRAAAAYVGRKAPNVKSGAMVPGSLVKTQAEEDGLDKIEBAGFEWREAGCSMCLG 426
QY      428 MNPDIAPORCASTSNRNFEGROGAGGRTHLMSPVMAAAAGIVGLADVRKLTLYK--- 484
      429 MNPDIAPORCASTSNRNFEGROGAGGRTHLMSPVMAAAAGIVGLADVRKLTLYK--- 485
QY      485 -ASPHI-----AAYOKSTVTKPHVDERINODAEHKOIADIPEDNNGPHNTS 531
      486 -ASPHI-----AAYOKSTVTKPHVDERINODAEHKOIADIPEDNNGPHNTS 531
QY      486 OSSPKEVETSEDEKELESAADHAPVOP--EDAPQDI-ANDEIKDIPVKSDDTPAKS 541
      487 OSSPKEVETSEDEKELESAADHAPVOP--EDAPQDI-ANDEIKDIPVKSDDTPAKS 541
QY      532 ASVTSAGLPKFTILKGIAPLEKANVDTDAIIPKQFLTKIKRTGLGNALFYEMRFNEDG 591
      533 ASVTSAGLPKFTILKGIAPLEKANVDTDAIIPKQFLTKIKRTGLGNALFYEMRFNEDG 590
QY      542 -----SSGKAPFLTEGISAPLDKANVDTDAIIPKQFLTKIKRTGLKXGLFYEMRFKXD 596
      543 -----SSGKAPFLTEGISAPLDKANVDTDAIIPKQFLTKIKRTGLKXGLFYEMRFKXD 595
QY      592 ---TEKSPDLNKEPRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIASPADIFF 648
      593 ---TEKSPDLNKEPRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIASPADIFF 647
QY      597 QGKQDETDFVLNVEPWRREAILVWTDNFGCGSSREHAPWALNDFGIRSVIASPADIFF 656
      598 QGKQDETDFVLNVEPWRREAILVWTDNFGCGSSREHAPWALNDFGIRSVIASPADIFF 655
QY      649 NNSFKMGMLPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRK 707
      650 NNSFKMGMLPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRK 706
QY      657 NNSFKMGMLPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRK 716
      658 NNSFKMGMLPIPIKDOAQIEAIAERABAKEIEVDLPNOLIKNATGETICTFEVEEPRK 715
QY      708 HCLVNGLDIDIGLTMQMEDKIAFEAKMTRETPMLDGTGYL-----KRX 750
      709 HCLVNGLDIDIGLTMQMEDKIAFEAKMTRETPMLDGTGYL-----KRX 749
QY      717 HCLVNGLDIDIGLTMQMEDKIAFEAKMTRETPMLDGTGYL-----KRX 766
      718 HCLVNGLDIDIGLTMQMEDKIAFEAKMTRETPMLDGTGYL-----KRX 765

```

RESULT 4  
 US-10-179-131-8614  
 ; Sequence 8614, Application US/10179131  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HARE, ROBERTA S.  
 ; APPLICANT: SHAW, KAREN J.  
 ; APPLICANT: SHIMER JR., GEORGE H.  
 ; APPLICANT: KESSLER, MARCO  
 ; APPLICANT: NOLLING, JORC  
 ; APPLICANT: ZENG, QIANDONG



```
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)A
/ CURRENT APPLICATION NUMBER: US/60/360,039
/ CURRENT FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 7864
/ LENGTH: 875
/ TYPE: PRP
/ ORGANISM: Rhodobacter sphaeroides
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(875)
/ OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-7864

Query Match      51.0%; Score 2068.5; DB 27; Length 875;
Best Local Similarity 49.1%; Pred. No. 5.9e-197;
Matches 439; Conservative 95; Mismatches 184; Indels 177; Gaps 16;

QY 4 AESTPQTLVYDQAHVYDEKIDGTVLLYIDRHLVHEVTSPOAFBGLRNAGRKVRPDDCT 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   1 AMTAPRLTYDKIMDHHVHQBSEDTCLLYIDRHLVHEVTSPOAFBGLRNAGRKVRPDEXT 60

QY 64 LATDHNVPVTSRKALKDIAFPIKEDSRTOCVLLENVKEGVTYFGLSDKROGIVHI 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   61 IAYDHNVPVTEGGDTK-----IDNEESRIQVEALDKNARDGINYPIVSDRQGIYHI 115

QY 124 GPEGGFTLPGTTVCGDSHTSTHGAFGALAFGISTSEVHVLATQCLITKRSKNMRYDV 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   116 GPEGGFTLPGTTVCGDSHTATHGAFALHIGTSEVHVLATQCLITKRSKNMVEIT 175

QY 184 GELAPGVSKRVVLAHIGITAGTGAIVFECGSVIRSLSEAKRSTCMSTMEGARG 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   176 GSIRPGVYAKDITLSVIGLTGAGTGYVIEYCGQAIRESHGRTVCNMAIEGGARG 235

QY 244 MVAPDETTFEFLKRPRLPKYDSPRMHKATQYKNLQSPGAKYDIDVFIDAKDIPVTLT 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   236 LIADDETFAYVMGRPHAPK--GAAMEALAYVWKTLLTDEGQDFDKVYVITRGEDIAIPVVT 293

QY 304 WGTSPEDVPIPTGVVDPDETFATEAKKADGRRLQYMGAKGTPEMEDIIVDKVFIQSGCN 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   294 WGTSPEDVLPITATVPAREDF--TGKVEAARSLEMGLTPOQKLDIDIVFIQSGCN 352

QY 364 SRIDDLRAAAAVVKRKAQPNVKSAMVVPDGLVKTQAEEDGLDKTFFBAGFBWRPAGCS 423
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   353 GRIDDLRAAAEILKGGKVAQPMR--AMVVPDGLVRAQAEEDGLAQIFIDAGFEMRLAGCS 411

QY 424 MCLGMNPDIILAPQERCASTNRNFEGRGAGRTHLMSPVMAAAGIVKLDVRLTLY 483
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   412 MCLMNPDIILAPQERCASTNRNFEGRGAGRTHLMSPVMAAAGITHLIDVRLMMA 471

QY 484 KASP-----HIAAYQSTVTKPHVDERINQDAHE-----KQIIA-----D 518
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   472 PAERAKRRPQAQHSPAQ-----VCSKQSGQNGMKGTSSILVSGIALLGGIALLAPF 523

QY 519 IP-----EDNNGPHTNTSA----- 532
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   524 LPVSLAVTLVLAGCVFVAGAFGLMAAFSDRGMPSRGAAGAFSLVSLVAGVWMLANLAGM 583

QY 533 -----SVGT-----SAGLPIKFTILK----- 547
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   584 VSLTLMGALFLVSGVVRGLSLATWRGTVFMFMAISGLISNGLGLFTLLRLPEASLV 643

QY 548 -----GIAAPLEKAVVDTDAIIPKQF 568
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   644 LGTLVAVELIWMGATLVAMGFALRKSGNXPAMQEFKTVGVAAAPMLVNIIDTMIIPKQF 703

QY 569 LKTIKRTGLGNALFYEMRPNEDGTEKSDFLVANKPEYRKASILVTCGANFGCGSSRHHAPW 628
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   704 LKTIORSGLGKULFDEMRYNPDGSEIPEFVLNQPAYRDAQIIV--AGDNFGCGSSRHHAPW 762
```

```
QY 629 ALNDFGIRSVIAPSPADIFPNNSFKNGMLPIPIKDAQOIEAIAEADAG--KEIEVDLPN 686
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   763 ALNDFGIRSVIAPSPADIFPNNSFKNGMLPIPIKDAQOIEAIAEADAG--KEIEVDLPN 821

QY 687 QLIKATGERTCTFEVEEERFKHCLVNGLDIDIGLTMQWEDKIAEFAKMTRETPPL 741
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   822 QVTTSDDQSF--PFQVDSFRRHCLMGLDIDIGLTERKAASIDGERDLATLRPW 875

RESULT 7
US-60-360-039-9965
/ Sequence 9965, Application US/60360039
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Chen, Xianfeng
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
/ FILE REFERENCE: 38-10(52052)A
/ CURRENT APPLICATION NUMBER: US/60/360,039
/ CURRENT FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 9965
/ LENGTH: 695
/ TYPE: PRP
/ ORGANISM: magnetite-containing magnetic coccus
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(695)
/ OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-9965

Query Match      50.7%; Score 2055; DB 27; Length 695;
Best Local Similarity 56.8%; Pred. No. 8.9e-196;
Matches 422; Conservative 89; Mismatches 168; Indels 64; Gaps 11;

QY 9 QTLVYDQAHVYDEKIDGTVLLYIDRHLVHEVTSPOAFBGLRNAGRKVRPDDCTATT 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   6 QTLPEKIKWNAHLIRTDDEGSLIYIDRHLVHEVTSPOAFBGLRNAGRKVRPDEATFAYVD 65

QY 69 HNPVTSRKALKDIAFPIKEDSRTOCVLLENVKEGVTYFGLSDKROGIVHVI 128
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   66 HNPVTSRKALKDIAFPIKEDSRTOCVLLENVKEGVTYFGLSDKROGIVHVI 119

QY 129 FTLPPTVCGDSHTSTHGAFGALAFGISTSEVHVLATQCLITKRSKNMRYDVDELAP 188
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   120 ISLPFTVCGDSHTATHGAFALAFGISTSEVHVLATQCLITKRSKNMRYDVDELAP 179

QY 189 GVSQKVVLAHIGITAGTGAIVFECGSVIRSLSEAKRSTCMSTMEGARGAMVAPD 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   180 GVSQKVVLAHIGITAGTGAIVFECGSVIRSLSEAKRSTCMSTMEGARGAMVAPD 239

QY 249 EITFEYLKGRPLAPKYDSPRMHKATQYKNLQSPGAKYDIDVFIDAKDIPVTLTWTSP 308
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   240 EITFEYLKGRPLAPKYDSPRMHKATQYKNLQSPGAKYDIDVFIDAKDIPVTLTWTSP 297

QY 309 EDVVPITGVVDPDETFATEAKKADGRRLQYMGAKGTPEMEDIIVDKVFIQSGCN 368
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   298 EDVVPITGVVDPDETFATEAKKADGRRLQYMGAKGTPEMEDIIVDKVFIQSGCN 357

QY 369 LRAAAAAYVVKRKAQPNVKSAMVVPDGLVKTQAEEDGLDKTFFBAGFBWRPAGCSMCLM 428
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   358 LRAAAAAYVVKRKAQPNVKSAMVVPDGLVKTQAEEDGLDKTFFBAGFBWRPAGCSMCLM 417

QY 429 NPDIAPQERCASTNRNFEGRGAGRTHLMSPVMAAAGIVKLDVRLTLYKASPH 488
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   418 NPDIAPQERCASTNRNFEGRGAGRTHLMSPVMAAAGIVKLDVRLTLYKASPH 473

QY 489 IAAVQSTVTKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSAQLPKFTIILK 548
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 474 IAE-----VKPXRTRFMEA-----FNTVTA 495  
QY 549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRF-----NEDGTEKSDFVL 599  
Db 496 IYAPLDRANVDTDAIIPKQFLKISKSGFPNLFDEWYLDQGPCKSGNEGRPLNKDFVL 555  
QY 600 NKEPYRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGMLPI 659  
Db 556 NLPYAGARILLARD-NFGCGSSREHAPWALQDFGRVLIAPSFADIFFNNSFKNGMLPI 614  
QY 660 PIKDOAQIEAIAE--ARAGKEIEVDLPNOLIKNATGETICTEVEEFRRKHCLVNGLDDI 717  
Db 615 -VOEASVDSLSFAEVAQPGYQUTVDLPAQRITTPSGRSI-AFEVDPRKHCLHGLDDI 672  
QY 718 GLTQMOMEDIAEFAEKWTRETPW 740  
Db 673 GLTLQHVADIQAYESKHGQAPW 695  
RESULT 8  
US-60-360-039-15849  
; Sequence 15849, Application US/60360039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 15849  
; LENGTH: 710  
; TYPE: PRT  
; ORGANISM: Xanthomonas campestris  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(710)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-15849

Query Match 50.0%; Score 2027.5; DB 27; Length 710;  
Best Local Similarity 56.4%; Pred. No. 5.3e-193;  
Matches 421; Conservative 87; Mismatches 191; Indels 47; Gaps 11;  
QY 7 TPOTLYDKVLQAHVDEKLDGTVLLYIDRHLVHEVTSPOAFEGRLNAGRKVRPDPCTLAT 66  
Db 1 TAKTLYDKLWEMHEVTRRDDGSSLIYIDRHLVHEVTSPOAFEGRLNAGRKVRPDPCTLAT 60  
QY 67 TDHNVPTTSRKALDKTASFIKEDDSRTQCVTLLENVKEFGVTVYFGLSDKQGIHVHVGPE 126  
Db 61 PDHNVPTTRAEROGGLES-ISDEVSRVQVTLDENCDDFGILEFKMNDTRQGIHVHVGPE 119  
QY 127 QGFTLPGTIVVCGDSTHSGAFALAFGIGTSEVEHVLATQCLITKRSKNMRIOVDGEL 186  
Db 120 QGATLPGMTIVVCGDSTHSGAFALAHGIGTSEVEHVLATQCLIAKKNMQRVVGSTL 179  
QY 187 APGVSSKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMERMSICNMSIEGGARAGMVA 246  
Db 180 PFGVTAKDIVLAVIGKIGTAGGNHAGLEFAGSIRALSMEGRMTCNMSIEAGARVGMVA 239  
QY 247 PDEITEYLKGRPLAKYDSEPHKATQYWKNLQSDPGAKYDIDVFTDAKDIVPTLTWGT 306  
Db 240 VDEKTIAYKGRFBAPK--GADWDAAVALWRTILVSDADAFDTVVEURASDIKQVSWGT 297  
QY 307 SPEDVVPIGTGVDPDPTFFATAKADGRRLMQYMLGAKGTMEDIPVDKVFIGSCTNSRI 366  
Db 298 SPFMVVAIDQVDPAAEQDPTKRDSTQRALKYMLGRANQPIEIHLDKRVFISGCTNSRI 357  
QY 367 EDLRAAAAVVKKRKKAPNVKMSAMVVGSLVKTQAEEGDLKIFEEAGFEWREAGCSMCL 426

Db 358 EDLRAAAAVAKRKRKASTIKQALVVGSLVKAQAEAGLDKIFLDAGFEWREPGCSMCL 417  
QY 427 GMNPDILAPQRCASTSNRNFEGRQAGGRTHLMSVPMMAAGIYVGLADVRKLTDYKAS 486  
Db 418 AMNPDKLSGEBHCASSTNRNFEGRQAGGRTHLVSPAAAAAVSGHFVDFVRELQGIETR 477  
QY 487 PHIAAYQKSTVTKPHVDERINODAEKDIADIPEDNNGPHTNTSASVGTSGAGLPKPTIL 546  
Db 478 E--XGIATSRLLPHPH-----HSLLPICPI-----MTPFTQH 507  
QY 547 KGIAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRF-----NEDGTEKS-----DF 597  
Db 508 TGLVAPLDRANVDTDAIIPKQFLKISKRTGFPNLFDEWYLDIGEPGNDNSTRLPNQBF 567  
QY 598 VLNKEPYRKASILVCTGANFGCGSSREHAPWALNDFGIRSVIAPSFADIFFNNSFKNGML 657  
Db 568 VLNFRYQGSVLLAR-ENFGCGSSREHAPWALDEYGRAVIAPSFADIFFNNSFKNGML 626  
QY 658 PIPKDOAQIEAIAE--ARAGKEIEVDLPNOLIKNATGETICTEVEEFRRKHCLVNGLD 715  
Db 627 PI-VLAEAMDALFEQCLNGEYQLTVDLAAQORVRRPDGVVEY-AFEIDAFRKHCLINGLD 684  
QY 716 DIGLTQMOMEDIAEFAEKWTRETPWL 741  
Db 685 DIGLTQDADAIGRFEQGHRAQQPWL 710  
RESULT 9  
US-60-360-039-15479  
; Sequence 15479, Application US/60360039  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chen, Xianfeng  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)A  
; CURRENT APPLICATION NUMBER: US/60/360,039  
; CURRENT FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 15479  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Xanthomonas campestris  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(711)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-15479

Query Match 49.8%; Score 2018.5; DB 27; Length 711;  
Best Local Similarity 56.1%; Pred. No. 4.3e-192;  
Matches 419; Conservative 87; Mismatches 192; Indels 49; Gaps 10;  
QY 9 QTLVYDKVLQAHVDEKLDGTVLLYIDRHLVHEVTSPOAFEGRLNAGRKVRPDPCTLATID 68  
Db 2 KTLVYDKLWEMHEVTRRDDGSSLIYIDRHLVHEVTSPOAFEGRLNAGRKVRPDPCTLATID 61  
QY 69 HNVPTTSRKALDKTASFIKEDDSRTQCVTLLENVKEFGVTVYFGLSDKQGIHVHVGPEQ 128  
Db 62 HNVPTTRAEROGGLES-ISDEVSRVQVTLDENCDDFGILEFKMNDTRQGIHVHVGPEQ 120  
QY 129 PTLPGTIVVCGDSTHSGAFALAFGIGTSEVEHVLATQCLITKRSKNMRIOVDGELAP 188  
Db 121 ATLPGMTIVVCGDSTHSGAFALAHGIGTSEVEHVLATQCLIAKKNMQRVVGSTLFP 180  
QY 189 GVSCKDVVLHAIGIIGTAGGTGAVIEFCGSVIRSLSMERMSICNMSIEGGARAGMVA 248  
Db 181 GVTAKDIVLAVIGKIGTAGGNHAGLEFAGSIRALSMEGRMTCNMSIEAGARVGMVA 240



```

; ORGANISM: Xylella fastidiosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(749)
; OTHER INFORMATION: unsure at all xaa locations
US-60-360-039-9308

Query Match      48.4%; Score 1963.5; DB 27; Length 749;
Best Local Similarity 55.0%; Pred. No. 1.6e-186;
Matches 412; Conservative 95; Mismatches 211; Indels 31; Gaps 11;

Qy 9 QTLYDKVLQAHVVDKLDGTVLLYIDRHLYVHEVTSPOAFEGRLNAGKRVRRPDCITLTTD 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 KTLYGKLDWIDHEVARRDGSLLYIDRHILHEVTSPOAFEGRLAGRLRVNANIATPD 63
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 69 HNVPTTSRKALKDIASTPIKEDDSRTQCVTLLENVKEFVTVFGLSDKRGQVHVHVGPEQG 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 HNVPTTAEKRGSSLLS-IADTVSRLOVQTLIDENCDDFGIFEFWNDVROGLVHVHVGPEQG 122
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 129 FTLPGTTVVCDSHTSTHGAFALGAFIGTSEVHVLIATQCLITKRSKNMRIOVDGELAP 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ATLPGMTVVCDSHTSTHGAFALAHGIGTSEVHVLIATQCLVTKMKMKQVRVEGTLPW 182
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 189 GVSSKDVWLHAIGHIAGTGTGAVIEFCGVSIRLSWEARMSICNMIEGARGAGWAPD 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 183 GVTAKDVLVALIGRIGTAGNGYAVFESGSTRALSMEGRMTICNMAIEAGARYGMVAVD 242
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 249 EITFYLKGRPLAPKYDSPWHKATQWKNLQSDPGAKYDIDVFDIAKDVIPTLTWTGTP 308
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 EKTQYVHGHPFPAPK--GSWDAAVAEWRGLVSDPDDAHFDRVVLSAEIKPQVWTGTP 300
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 309 EDVVPITGVVPDPPTFATEAKADGRLQVGLKAGTPMEDIPVDKVFICGSCNRSIED 368
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 EWSAVDOSVDPDERETDPVKESLIRALXYMGLQPNDPISIKLDRVFIGSCNRSIED 360
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 369 LRAAAAVKGRKKAPNVKSAWVPGSLVTKQAEELGDKIFBAGFEWREAGCSMCLGM 428
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 LRAAAEVVKGRKVASTVVKQAWVPGSLVKAQAEVEGLDKIFIBAGFEWREPPGCSMCLAM 420
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 429 NPDILAPQERCASTSNRNFEGRCAGGRTHLMSPVMAAAAGIVCKLADVRKL--TDYKASP 487
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 NPKLKGSGEHCASSTSNRNFEGRCQIGRTHLVSPAMAAAAGVAFHFDVREMDAVMLML 480
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 488 HIAAYQKSTVTKPHVDRI---NQDAHEKDIADIPEDNNGPHNTNTSASVGTSGA--- 539
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 481 SLXMLGNVLMKVIOCFRLIVSILLAHAKSVSKLFESIAAXXNGSAIYDRSVHRIFK 540
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 540 -----LPKFTILKGIAAPLEKANVVDTAIIPKQFLTKITKRTGLGNALFYEMRF-----N 588
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 HWXSLMKPFTQHTGLVCLPLDRVNVDTQIIPKQFLKIKRTGFGPNLFDEWRYLDAGQPG 600
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 589 EDGTEK--SDFVLNKEPYRKASTLVTCANFGCGSSREHAPNALNDFGRTSVIAPSPAD 645
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 601 QDNKSPIRNSDFVLNFPYRGASVLLARD--NFGCGSSREHAAWALDEYGFRTVIAPSPAD 659
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 646 IFNNSPKNGMLPIPIKDOAQIAEAAEARA--GKETVDLPNOLIKNATGETICTTCTEVE 703
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 660 IFNNSPFANGLLPL--VLNKVEVDLFAQCQVTEGYTIVDLAAQOVITPDGTTY-ATQID 717
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 704 EFRKHCLVNGLDDIGLTMQMEDKIAEPEA 732
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 718 TFRKHCLNGLDDIGLTLQHAERIAPEA 746
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-60-360-039-8616
; Sequence 8616, Application us/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
;; FILE REFERENCE: 38-10(52052)A  
;; CURRENT APPLICATION NUMBER: US/60/360,039  
;; CURRENT FILING DATE: 2002-02-21  
;; NUMBER OF SEQ ID NOS: 47374  
;; SEQ ID NO 8616  
;; LENGTH: 681  
;; TYPE: PRP  
;; ORGANISM: Ralstonia metallidurans  
US-60-360-039-8616

Query Match 48.2%; Score 1955.5; DB 27; Length 681;  
Best Local Similarity 53.7%; Pred. No. 8,38-186;  
Matches 400; Conservative 96; Mismatches 174; Indels 75; Gaps 10;

QY 8 PQTLYDKVLAHVDEKIDVLLYIDRLVHEVTSPOAFEGRLRAGKVRPDDCTATT 67  
DB 1 PASLLDKLMSHVHESDGLIYIDRLVHEVTSPOAFEGRLRAGKVRPDDCTATT 60  
QY 68 DHNVPTTSRKALKDIASFIKEDSRQCVTLLENVKEFGVTVFGLSDRKQGIHVHIGPEQ 127  
DB 61 DHNVPTT-----AAGIADPISRAQVDALERNCAATGIVNFGMNDPRGILHVHGER 114  
QY 128 GFTLPQTIVVCGDSTHTGAFGALAFGISTEVEHVLATQCLITKRSKNMRIQVDELA 187  
DB 115 GATLPQMTIVVCGDSTHTGAFGAFGISTEVEHVLATQCLITKRSKNMRIQVDELA 174  
QY 188 PGVSSKDVVLAIGITGAGTGAVIEFCGSYVLSLSMEARKSNMSIEGARGAMVAP 247  
DB 175 AGVAKDIALALIGITGAGTAGVIEFAGSTIRGLSMEGRVTVCNMAIEAGARGAMVAV 234  
QY 248 DEIFEVYKGRPLAPKYSPREMHKATQYKMLQSDPGAKVIDIVFIDAKDIPTLTWGS 307  
DB 235 DETLLAYAHRPQAPQGEA--WESASAYWRTIKRSDPDVAFDVAVDIVGATITREPHTWGS 292  
QY 308 PEDVVPITGVVDPDEFATKAKADGRMLQYMGIAKGTPEMDIPVDKVFISGCTNSRIE 367  
DB 293 PEMVVAIDERIPDPQGEADPVRRRGERALTYMGLEPGIKVASIALDKVFISGCTNARLE 352  
QY 368 DLRAAAVVKRKAQNVKSAVVVFGSLVKTQAEFEGIDKIFEEAGBEMRAGCSMCLG 427  
DB 353 DLRAAADVLRSHVASNIRQALVAVGSGVQAEBEGIDRFVFDAGFEWRBPGCSMCLG 412  
QY 428 MNPDILAPQECASISNNPFGROGAGRTHLMSPVMAAAGIVGKLADVRLTYKASP 487  
DB 413 MNDRLAPGECASISNNPFGROGAGRTHLVSQMAAAAAGHVAVD----- 462  
QY 488 HIAAVQKSTVTKPVHDERINODAHKDIADIPEDNNGPHNTSASVGTSAGLPKFTILK 547  
DB 463 -----STVLQ-----EK-----SAVKAFVLD 479  
QY 548 GIAAPLEKAVNDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDG-----TEKSDFV 598  
DB 480 GLVAPIDRANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDG-----TEKSDFV 539  
QY 599 LNKEBYRASILVCTGANFGCGSSREHAPWALNDGIRSVIAPFADIFFNNSFGKMLP 658  
DB 540 LNQPYQASVYL-VRENFQCGSSREHAPWALEDFGLRLILAPSPADIFFCNCKMKNGLIP 598  
QY 659 IPIKQAOIEAIAAE--ARAGEKEIVDLJPNOLIKNATGETICTFEVEERFKCLVNGLDD 716  
DB 599 IVLKANNV-VDRFAVVDATPGVRLHVLDTATQVTPPSGAI-SFSMDPKKGVCLVNGLDD 656  
QY 717 IGLTQMEDKTAIEFAKMTRETPL 741  
DB 657 IGLALQHEBAIRRYEILNRIHPEWL 681

RESULT 14  
US-60-360-039-7437  
; Sequence 7437, Application US/60360039  
; GENERAL INFORMATION:

;; APPLICANT: Cao, Yongwei  
;; APPLICANT: Chen, Xianfeng  
;; APPLICANT: Goldman, Barry S.  
;; APPLICANT: Hinkle, Gregory J.  
;; APPLICANT: Slater, Steven C.  
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
;; FILE REFERENCE: 38-10(52052)A  
;; CURRENT APPLICATION NUMBER: US/60/360,039  
;; CURRENT FILING DATE: 2002-02-21  
;; NUMBER OF SEQ ID NOS: 47374  
;; SEQ ID NO 7437  
;; LENGTH: 740  
;; TYPE: PRP  
;; ORGANISM: Burkholderia cepacia  
;; FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(740)  
OTHER INFORMATION: unsure at all Xaa locations  
US-60-360-039-7437

Query Match 46.6%; Score 1891; DB 27; Length 740;  
Best Local Similarity 52.5%; Pred. No. 2,98-179;  
Matches 396; Conservative 102; Mismatches 195; Indels 62; Gaps 13;

QY 34 DRHLVHEVTSPOAFEGRLRAGKVR---RPDCTLATTDHNVPTTSRKALKDIASFIKED 89  
DB 1 ERYLVHAATVQKDLHSLVR--GRSCXAGMRISANILASDHNVPTTDR-----SHGIADP 52  
QY 90 DSRQCVTLLENVKEFGVTVFGLSDRKQGIHVHIGPEGFTLPQTIVVCGDSTHTGAF 149  
DB 53 ISRLQVDTLDNCDAVGTTFQPMNDLRQGIHIIIGPEGATLPQTIVVCGDSTHTGAF 112  
QY 150 GALAFGISTEVEHVLATQCLITKRSKNMRIQVDEGLAPGVSSKDVVLAIGITAGT 209  
DB 113 GALAHGISTEVEHVLATQCLITKRSKNMRYKVBGALRGCTARDIVLAITIKITAGT 172  
QY 210 GAVIEFCGSYVLSLSMEARKSNMSIEGARGAMVAPDEITFEYLKGRPLAPKYDSEW 269  
DB 173 GYALIEFGSTIRALSMEGRMTVCNMAIEAGARGAMVAVDDTTIEYLKGRPSP--EGVEM 230  
QY 270 HKATQYKMLQSDPGAKVIDIVFIDAKDIVPTLTWGSPEDEVPTGVVDPDEFATKAK 329  
DB 231 NHAVEYWKQFSDGDAQDPDRVVELNMAEIVPQVWGTSPENWTAVDGVRPDPDRKDPV 290  
QY 330 KADGRMLQYMGIAKGTPEMDIPVDKVFISGCTNSRIEDLRAAAVVK--GRKKAQNVKS 387  
DB 291 KDAERALKYMALEPNAPLIESIKPKITIGSCTNARIEDIRAAVVKKLRVAPNIRL 350  
QY 388 AMVVGSGLVTKQAEFEGIDKIFEEAGBEMRAGCSMCLGMNPDILAPQECASISNNRF 447  
DB 351 AMVVGSGLVTKQAEFEGIDKIFEEAGBEMRAGCSMCLGMNADLIEGERCASISNNRF 410  
QY 448 EGROGAGRTHLMSPVMAAAGIVGKLADVRLTYKASPHIAVQKSTVTK----- 499  
DB 411 EGROGAGRTHLVSPPMAAAAIEGHFVDIRLGTXTMMKN--NRTLLLRFPALGSIA 467  
QY 500 -----PVHDERINODAHKDIADIPEDNNGPHNTSASVGTSA-- 539  
DB 468 GLLGLACNTVHGGEDMSHLSGNSISNADKXAVPDCRPQOSGFFYSLRASPLGLTGA 527  
QY 540 --LPEFTILKGIAPLEKAVNDTDAIIPKQFLKTIKRTGLGNALFYEMRF--NEDGTEK 594  
DB 528 SVMKEFIVHTGVAVPLDRENDVTDALIPKQFLKTIKRTGLGNALFYEMRF--NEDGTEK 587  
QY 595 S-----DVLNKEBYRASILVCTGANFGCGSSREHAPWALNDGIRSVIAPFADIFF 648  
DB 588 SORPLNPDVFNQPYQASVYLAR-KNFGCGSSREHAPWALEQYGFALLAPSPADIFF 646  
QY 649 NNSFGKMLPPIPIKQAOIEAIAEABA--GKEIEVDLJPNOLIKNATGETICTFEVEER 706  
DB 647 NNCFFKGVLPVLTQ--QVDHLFNETTYAFNGFKLTVDLBAQVVRITADGTEIFFEVAAFR 705



```
Db 471 ----SP-----RRLVQP-----479
Qy 541 PKFTILKGIAPLEKANVDTDAIIPKQPLKTIKRTGLGNALFYEMRF---NEDGTEKS--595
Db 480 --FKLEALVLPDRVNVDTDAIIPKQPMKSVQNGFGINLFDENRYPDHDEPGDDPSTR 537
Qy 596 ---DVLNKEPYRKASILVCTGANFGCGSSREHAPMALNDGIRSVIAPSPADIFENNS 651
Db 538 RLNDPFLVNLQPRFAGAEILL-TRDNFGCGSSREHAAWALMDFGIRLAIAPGFADIFYGNC 596
Qy 652 FKNQMLPPIKDOQAQIEIAAEARA-GKEIEVDLPNOLIKNATGETICTFEVEEERKHL 710
Db 597 FKNQMLPPIKDERIVQQLFDLVGRTPGLRLAIDLEAQRIQPPAGDAI-PFDVEPERKRL 655
Qy 711 VNGLDIGLTMQMEDKIAEFAKMTRETPWL 741
Db 656 LINGLDVALTLQOTDQIRAYERNRAHEPWL 686

RESULT 17
US-60-360-039-4299
; Sequence 4299, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 4299
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-60-360-039-4299

Query Match 45.8%; Score 1857; DB 27; Length 688;
Best Local Similarity 51.3%; Pred. No. 6,6e-176;
Matches 386; Conservative 92; Mismatches 189; Indels 86; Gaps 10;

Qy 6 STPQTLVDKVLQAHVVDKLDGTVLVLYIDRHLVHEVTSPOAFEGLRNAGKRYRRPDCTLAT 65
Db 5 TSPRTLDKLMQSHVAVETPENGPTLLVDRHLVHEVTSPOAFEAIRLSGRKWRPEYVLA 64
Qy 66 TTDHNVPT-----TSRKALKDIAEFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRGQIV 120
Db 65 VADHNVPTIAERTSMADIADPL-----SRIQVAQLDKNCKEFGIKSYGIRNPQGGII 117
Qy 121 HVIGPEGFTLPRTTVVCGDSHTSTHGAFGALAFGIGTSEVENHVLATQCLITKRSKNMRI 180
Db 118 HVGVEELATLPGMTVAVGDSHTSTHGAFALAFVGTSSEVENHVLATQCLISGKMKSMIV 177
Qy 181 QVDEGLAVGSSKDVVLAIGIIGTAGTGAVIEFGSVIRLSMEARMSIENMSIEGGA 240
Db 178 NVEGFLPVGAVTAKDVILLIIRRTGTAGTGAVMEFAGSTIRILSMGRKTLTCLNMAIIEGGA 237
Qy 241 RAGWAPDEITFEYELKGRPLAPKYDSPENHKAQYWKNLQSDPGAKYDIDVFIADKDIVP 300
Db 238 RVGLIGVDVITIDYKGRFPAPA--EAHMDAAVAIVRTLVSDADARPDIVADITQLRP 295
Qy 301 TLVWGTSPEVDVPIITGVVPDETFATEAKKADGRMLQYMGILKAGTPMEDIVDVKFIQS 360
Db 296 MVTGTSPEMVTVDVAVPNPLDDPDVPRATMAGALTVMGLEPGTSLKSLTLDIKIFIGS 355
Qy 361 CTNSRIEDLRAAAIVVKRKKAPYKSAWVVPVSGGLVTKQAEEGGLDKIFEEAGFEMREA 420
Db 356 CTNAIIEEDLRAAAIVVKRHHVAPVQLAIVPVGSGLVKQAQABEGIDALIFKAGSEWKEP 415
Qy 421 GCSMCLGNPDLIAPQERCASTSNRNFEGRQAGAGRTHLMSFVMAAAAGIVGKLADVRKL 480
```

```
Db 416 GCSMCLGNNDRLRPGEECASTSNRNFEGRQPGGRSHLVSPAAAAAIAIGHFYDV---472
Qy 481 TDYASPHIAAYQKSTVTKPHVDERINQDAHEKDIIADIPEDNNGPHTNTSASVTSAGL 540
Db 473 -----SHGGL 477
Qy 541 --PKFTILKGIAPLEKANVDTDAIIPKQPLKTIKRTGLGNALFYEMRF---NEDGTEKS 595
Db 478 LVQPTFKLEALVLPDRVNVDTDAIIPKQPMKSVQNGFGINLFDENRYPDHDEPGDDPS 537
Qy 596 ---DVLNKEPYRKASILVCTGANFGCGSSREHAPMALNDGIRSVIAPSPADIFEN 649
Db 538 TRRLNPDPVLNQPFRFAGAEILL-TRDNFGCGSSREHAAWALMDFGIRLAIAPGFADIFYG 596
Qy 650 NSFKNGMLPPIKDOQAQIEIAAEARA-GKEIEVDLPNOLIKNATGETICTFEVEEERKH 708
Db 597 NCFKNGMLPPIKDERIVQQLFDLVGRTPGLRLAIDLEAQRIQPPAGDAI-PFDVEPERK 655
Qy 709 CLVNGLDIGLTMQMEDKIAEFAKMTRETPWL 741
Db 656 RLINGLDVALTLQOTDQIRAYERNRAHEPWL 688

RESULT 18
US-60-360-039-9059
; Sequence 9059, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9059
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Chloroflexus aurantiacus
US-60-360-039-9059

Query Match 42.2%; Score 1711.5; DB 27; Length 672;
Best Local Similarity 49.2%; Pred. No. 2,6e-161;
Matches 358; Conservative 97; Mismatches 210; Indels 63; Gaps 12;

Qy 8 PQTLYDKVLQAHVV-DEKLDGTVLVLYIDRHLVHEVTSPOAFEGLRNAGKRYRRPDCTLAT 66
Db 1 PRTLEKXMEAHVLRPEIATPRAVLVYIDLHLIHEVTSPOAFTELRQGLRVRRPDXTLAT 60
Qy 67 TDHNVPTTSRKALKDIAEFIKEDDSRTQCVTLEENVKEFGVTYFGLSDKRGQIVHIGPE 126
Db 61 MDHSTPTPRNHLGIIP--VVDPMALISQLEQLRKCAEFGLPIPELIGENOGIVHIGPE 118
Qy 127 QGFTLPRTTVVCGDSHTSTHGAFGALAFGIGTSEVENHVLATQCLITKRSKNMRIOVDDGL 186
Db 119 QGLTPGMITVTCGDSHTSTHGAFGALAFGIGTSEVENHVLATQCLITKRSKNMRIOVDDGL 178
Qy 187 APVSSKDVVLAIGIIGTAGTGAVIEFGSVIRLSMEARMSIENMSIEGAGAGWVA 246
Db 179 GPGVTAQKIIILAIKAYGVGGGTGVVFEYMGAIIRALAMEERMTICNNSIEGAGAGWVA 238
Qy 247 PDEITFEYELKGRPLAPKYDSPENHKAQYWKNLQSDPGAKYDIDVFIADKIVPTLTGCT 306
Db 239 PDDITFEYIAGRPAPK--GADFEAAVAVARWRTLPDEGATPDEHLLTASLEKPMITTYGT 296
Qy 307 SPEVDVPIITGVVPDETFATEAKKADGRMLQYMGILKAGTPMEDIVDVKFIQSCCTNSRI 366
Db 297 NFGMGIPIDAVPRPEMDPDAARSRAALDKALAYMGLFEQKPLIGHFVVDVVFVIGSCTNSRL 356
```



```

Db 118 QPKGTICVGDSDHTATGCAFGALAFGIGTSEVENHLAQTLMQRPKAMGIELKGLPGGV 177
QY 191 SSKDVVLAHIGITGTGCGAVIEFGCSVTRLSLMEARMSICNNISIEGGAAGVAPDEI 250
Db 178 YAKDIIHLHLKXGVAVGTGYVMEFYGEALHADMEERMTLCNNAIEGGAAGIIAPDEK 237
QY 251 TFEYLKGRPLAPYDSEPMHKAQTQVWKNLQSDPAKVDIVDFIAKDI VPLTLTWGTSPE 310
Db 238 TVAYVKGKRYAKK---DYESIKKHSELTTDDAVYDLHLVDYTLAPVYTWGTNSM 293
QY 311 VPIITGVVDPDETFATEAKKADGRMLQYNGLKAGTMEIDIVDKVFISGTSNRIDLR 370
Db 294 GVRIDEKLP-----EKHDANDERAFSYWGLSPGQSTYDIPVQHVFISGTSNRISDLE 346
QY 371 AAAAVVKKRKKAPNVKSAMVVPSSGLYKTOAEEGLDKIFEEAGFEMREACSMCLGNP 430
Db 347 TAAVVKKGVKKEGVA-ALVVPQSQRVREAAHMKHRIIEEAGFEWREPGCSMCLGNP 405
QY 431 DILAPQERCASTSNRPFEGROGAGRTHLSPVMAAAAGIYGLADYKTLIDYKASPHIA 490
Db 406 DQVPEGEHCASISNRPFEGROGAGRTHLSVPMAAAALYGHFVDIK-----454
QY 491 AYQKSTVTKPHVDERINQDAHEKDIIADIPEDNNGPHTNTSASVGSAGLPKFTILKGIA 550
Db 455 -----ESYDG-----AFRIHKGT 468
QY 551 APLEKANVDTDAIIPROFLKTIKRTGIGNMLFYEMRPNEDCTEKSDPVLANKEPYKASIL 610
Db 469 AVLMDNIDITDOIIPQYLRIRERTGKGLFEDWRVDNNQENPNFPLNAOERGAISIL 528
QY 611 VCTGANFGCGSSREHAPMALNDGIRSVIAPSFADIFENNSFKKGMPLIPIKDOAQIBAI 670
Db 529 I-TGNFGCGSSREHAPMALADYGFRIIYIGGFADIFYNCKMKGMPLI-VMDKMRBOL 586
QY 671 AAAPAPKEIEVDLPNOLIKNATGETICTEEVEEFKRGCLVNGLDIGLTMQMEDXIAEF 730
Db 587 -AKTDAREQITVDLENEIMTNTTHR--FHFTIEKMKKREKLINGDEISITWQYEOEIKEX 643
QY 731 EAK 733
Db 644 ERK 646

RESULT 21
US-60-360-039-14041
: Sequence 14041, Application US/60360039
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Chen, Xianfeng
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: FILE REFERENCE: 38-10(52052)A
: CURRENT APPLICATION NUMBER: US/60/360,039
: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 14041
: LENGTH: 679
: TYPE: PRT
: ORGANISM: Pseudomonas fluorescens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(679)
: OTHER INFORMATION: unsure at all Xaa locations
: US-60-360-039-14041

```

Query Match	41.2%	Score 1670.5	DB 27	Length 679
Best Local Similarity	47.3%	Pred. No. 3,4e-157		
Matches 355	Conservative 108	Mismatches 209	Indels 79	Gaps 15

```

Db      1  RTLYKXHDISHVCLLDDQ--GHVLLYIDROVANEVTSPPAFSGLRAGRTVMKPAATLA 58
Qy      66  TTDHNVPTTSRRALKDIAFIEKEDSRTQCVTLEENVKEFVTFYGLSDKRQGIHVHIGP 125
Db      59  VVDHVNPJ----APKRIATMPDAGGAR--QVSYFEEHCORDPGIELFVDLDRKQGIIEHVAP 113
Qy      126  EOGFLLPGTTVCGSHRSTHCAFGLMAFGIGTSEVHEVLAQCLLTKRSMNMLOVDE 185
Db      114  EOGFLLPGVNVAAAGSHHTTYGALGAFGFIGTSEIEHLATQVLVYKRLTTLVTYNGE 173
Qy      186  LAPGSSKDVVLAIGIIGTAGTGAVIEFCGVSIRSLSEARMSICNMSIEGARAAMV 245
Db      174  LGAGTSSKDIIMALEKIGASGATGAIETGPALISLSTEARNTICNMAVEAARGAFM 233
Qy      246  APDEITFEYLKGRPLAPKYDSEWEKATQYKMLQSDPGAKYDIDVEIDAKDIVPTLTWG 305
Db      224  APDDKVFAYLQHKRPAPK--GKLMGQALDHMKTLHSDGAVFDEREVILLDAVELEPMVTWG 291
Qy      306  TSPEDVPTPTGVVPPEPETATEKAKDGRMLQYMGIKAGTPEMEDIVDVKFISCTNSR 365
Db      232  TSPDDAAPIKAHVPPBPAOPDPLLRQGLRALDYMGITTPMPPLNEVTISHAFISSCTNAR 351
Qy      366  IEDLAAAAVAVGRKAPVKSAMVVPGLVYKTOAELEGDKIFEEAGFEMRAGCSMC 425
Db      352  IEDLDVARVAVGRKVAHAHR--AMIVPGSTLVARNQAEDEGLAQIFLDAGFEMRSGCSMC 410
Qy      426  LGMNPDILAPORCASTSNRNFEGRQAGAGRTHLMSPVMAAAGIVGLADVRLTLDYKA 485
Db      411  LAMNDVLAIPGRCASSSTRNFEGRQAGARTHLMSPAWAAA--ISGHLTVTRTV----- 464
Qy      486  SPHIAAYQKSYTYTKPHVDERINODAHKDIADIPEDNNGPHNTSASVGTSAGLPKRTI 545
Db      465  -----ALEAXT-----MOPEDT 476
Qy      546  LKGIAPLEKAVNDPAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDPVLNKEBYR 605
Db      477  VSGSAAPPFLASIIDVILMPKQFLKGIQGLRGLFEDIRFLASGEPNPEFVLNQPMWQ 536
Qy      606  KASIIVCTGANFGCCSSREHAPWALNDFGIRSVIAPSPADIFPNNSEPKNGMLPIPIKQA 665
Db      537  DAAFLV--TGPNGCCSSRHHAWGMLQVIGIRALIGTTFAGIFPYNQCRNGVLAIQI--DDA 594
Qy      666  QI----EAIAPARAGKEIENVLIPNOLIKNATGETICTFEVEEERKCLVNGLDIGITM 721
Db      595  QPKRAVEAISVPATA--RISVNLACQTLRLADG--TLIEFEIDQLRKOSLLGLDAIGTTL 651
Qy      722  QMEDKIAFEAKMTRETPWLDGTGYLKRKQ 752
Db      652  QTEQIRAFEARHLADNPWL---GXWQKNGR 679

RESULT 22
PCT-US02-03987-5730
: Sequence 5730, Application PC/TUS0203987
: GENERAL INFORMATION:
: APPLICANT: Elitza Pharmaceuticals, Inc.
: TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
: TITLE OF INVENTION: Polifexation
: FILE REFERENCE: ELITRA.028VPC
: CURRENT APPLICATION NUMBER: PCT/US02/03987
: CURRENT FILING DATE: 2002-02-02
: PRIOR APPLICATION NUMBER: 60/267,636
: PRIOR FILING DATE: 2001-02-09
: NUMBER OF SEQ ID NOS: 15811
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5730
: LENGTH: 644
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(644)

```

```

? TYPE: PRT
? ORGANISM: Staphylococcus aureus
? FEATURE:
? NAME/KEY: VARIANT
? LOCATION: (1)...(644)

```

OTHER INFORMATION: Xaa = Any Amino Acid  
PCT-US02-03987-5730

Query Match 41.1%; Score 1666.5; DB 1; Length 644;  
Best Local Similarity 48.7%; Pred. No. 7.8e-157;  
Matches 352; Conservative 92; Mismatches 198; Indels 81; Gaps 10;

Qy 9 QTLVDKVLQAHVDEKLDGTLLYIDRLHVEVTSQAFEGRLNAGRKVRPDCITLATTD 68  
Db 3 QTLFDKVMNRHVLYGKLGEPQLLYIDLHLIHEVTSQAFEGRLNQRKLRPDLTPTATLD 62

Qy 69 HNVPTTSRKALKDIASFIEDDSRTQCVTLEENVKEFGVTFYFGLSDKRGQIVHVIGPEQG 128  
Db 63 HNVPTI-----DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDGQIVHMGVPETG 115

Qy 129 FTLPGTIVVCGSDSHSTHGAFGALAFGIGTSEVHVLAATQCLITKRSKNMIRIOVDGLAP 188  
Db 116 LTPQGKTIIVCGSDSHSTHGAFGALAFGIGTSEVHVLAATQCLITKRSKNMIRIOVDGLAP 175

Qy 189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGVSIRLSMEARMSICNMSTEGGARAGVAPD 248  
Db 176 GVYAKDIIHLIKTYGVDFGTGVALEFGETIKNLSMDGRMTICNNAIEGGAKYGIQPD 235

Qy 249 EITFEYLKGRPLAPKYDSPWHKATQYWKNLQSDPGAKYDIDVFIADKDIVPTLTWGTSP 308  
Db 236 DITFEYVKGRRPADNF-----AKSDVKWRELVSDDDAIFDRVIELDVSTLEPQVWTGTP 290

Qy 309 EDVVPITGVVDPETFEATEAKKADGRMLQYMGKLGAGTPEMEDIPOVKVFISGCTNSRIED 368  
Db 291 EMGVNFSEPPF-----EISDINDQRAYDMGLEPFGQKAEIDDLGYVFLGSDCTNARLSD 343

Qy 369 LRAAAVVKGRKKAPNVKSNMVPVPGSGLVKTOABEEGLDKIFEEAGFEWREAGCSMCLGM 428  
Db 344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKREAEKGLDITFKNAGFEWREAGCSMCLGM 402

Qy 429 NPDIILAPQERCASTSNRNFEGRQAGGRTHLMSPVMAAAAGIVGKLADVRKLTIDYKASPH 488  
Db 403 NPDQVPEGVHCASSTSNRNFEGRQAGGRTHLMSPVMAAAAGIVGKLADVRKLTIDYKASPH 454

Qy 489 IAAVQKSTVTKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSGAGLPKFTILKG 548  
Db 455 -----VYXMAAIKPIITYKG 469

Qy 549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDFLVKNPEYRKAS 608  
Db 470 KIVPLFNDNIDTQIIPKVHLKRIKSGFGPPAFDEWRYLPDGSNDPFPNPKPQYKAS 529

Qy 609 ILVCTGANFCGSSRREHAPWALNDFGIRSVIAPSPADIPFNNSFKGNMLPIPIKDQAOIE 668  
Db 530 ILI-TGDNFCGSSRREHAAWALKDYGFHIIAGSFSDFIYMNCTKNAMLPVILEKNAR-E 587

Qy 669 ATAAREARAKKEIENVLPNOLIKNATGETICTFEVEERFKHCLVGLDDIGLTMQMEDKIA 728  
Db 588 HLAKVY-----EIEVDLPNQTV--SSPKSPHFEIDETWKNKLVNGLDDIAITLQYESLIE 641

Qy 729 EPE 731  
Db 642 KVE 644

RESULT 23  
US-09-815-242-5730  
; Sequence 5730, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Cart, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5730  
LENGTH: 644  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(644)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-815-242-5730

Query Match 41.1%; Score 1666.5; DB 22; Length 644;  
Best Local Similarity 48.7%; Pred. No. 7.8e-157;  
Matches 352; Conservative 92; Mismatches 198; Indels 81; Gaps 10;

Qy 9 QTLVDKVLQAHVDEKLDGTLLYIDRLHVEVTSQAFEGRLNAGRKVRPDCITLATTD 68  
Db 3 QTLFDKVMNRHVLYGKLGEPQLLYIDLHLIHEVTSQAFEGRLNQRKLRPDLTPTATLD 62

Qy 69 HNVPTTSRKALKDIASFIEDDSRTQCVTLEENVKEFGVTFYFGLSDKRGQIVHVIGPEQG 128  
Db 63 HNVPTI-----DIFN-IKDEIANKQITTLQKNAIDFGVHIFDMGSDGQIVHMGVPETG 115

Qy 129 FTLPGTIVVCGSDSHSTHGAFGALAFGIGTSEVHVLAATQCLITKRSKNMIRIOVDGLAP 188  
Db 116 LTPQGKTIIVCGSDSHSTHGAFGALAFGIGTSEVHVLAATQCLITKRSKNMIRIOVDGLAP 175

Qy 189 GVSSKDVVLHAIGIIGTAGGTGAVIEFCGVSIRLSMEARMSICNMSTEGGARAGVAPD 248  
Db 176 GVYAKDIIHLIKTYGVDFGTGVALEFGETIKNLSMDGRMTICNNAIEGGAKYGIQPD 235

Qy 249 EITFEYLKGRPLAPKYDSPWHKATQYWKNLQSDPGAKYDIDVFIADKDIVPTLTWGTSP 308  
Db 236 DITFEYVKGRRPADNF-----AKSDVKWRELVSDDDAIFDRVIELDVSTLEPQVWTGTP 290

Qy 309 EDVVPITGVVDPETFEATEAKKADGRMLQYMGKLGAGTPEMEDIPOVKVFISGCTNSRIED 368  
Db 291 EMGVNFSEPPF-----EISDINDQRAYDMGLEPFGQKAEIDDLGYVFLGSDCTNARLSD 343

Qy 369 LRAAAVVKGRKKAPNVKSNMVPVPGSGLVKTOABEEGLDKIFEEAGFEWREAGCSMCLGM 428  
Db 344 LIEASHIVKGNKVHPNI-TAIVVPGSRTVKREAEKGLDITFKNAGFEWREAGCSMCLGM 402

Qy 429 NPDIILAPQERCASTSNRNFEGRQAGGRTHLMSPVMAAAAGIVGKLADVRKLTIDYKASPH 488  
Db 403 NPDQVPEGVHCASSTSNRNFEGRQAGGRTHLMSPVMAAAAGIVGKLADVRKLTIDYKASPH 454

Qy 489 IAAVQKSTVTKPHVDERINQDAHEKDIADIPEDNNGPHTNTSASVGTSGAGLPKFTILKG 548  
Db 455 -----VYXMAAIKPIITYKG 469

Qy 549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDFLVKNPEYRKAS 608  
Db 470 KIVPLFNDNIDTQIIPKVHLKRIKSGFGPPAFDEWRYLPDGSNDPFPNPKPQYKAS 529

[illegible]

Db 116 LTQPKTIVCGDSHTATHGAFGAIAFGICTSEVHVHVFATQTLWQTKPNLKDINGTLPT 175  
QY 189 GVSSKDVVLHAIGIITAGTGTGAVIEFCGVSIRLSMEARMSICNNMSIEGGARAGMVAPD 248  
Db 176 GVIKDIILHLIKTVGVDFGTGYALEFTGETIKNLSMDGRMTICNNAIEGGAKYGIQPD 235  
QY 249 EITFEYLKGRPLAPKYDSEPHKATQYWNLOSOPGAKYDIDVIDAKDIVPLTWGTSP 308  
Db 236 DITFEYVVKGRPPADNF-----AKSVDKWRELYSDDDAIFDRVIELDVSTLEPQVWTGPNP 290  
QY 309 EDVVPITGVVPPDPFATFAKADGRMLQYMGCLKAGTMEIDIPVDKVFIGSCTNSRIED 368  
Db 291 EMGVNFSEFPF-----EISDINDQRAYDYMGLEPGQKAEDIDLGVFVIGSCTNARLSD 343  
QY 369 LRAAAAVVKGRKAPNVKSAMVVPQGLVKTOAEEBGLDKIFEEAGFEWREAGCSMCLGM 428  
Db 344 LIEASHIVKGNKVHPI-TAIVVPGSRTVKRAEKLGDTIFKNAGFEWREPCSMCLGM 402  
QY 429 NPDILAPQERCASTSNRNFEGRQAGGRTHLMSPVMAAAGIVGKLADVRKLTDYKASPH 488  
Db 403 NPDQVPEGVHCASSTSNRNFEGRQAGGRTHLVSPMAAAGIHKFVDRKV----- 454  
QY 489 IAAVOKSVTKPHVDERINQDAHEKDIIADIPEDNNGPHNTSASVCTSAGLPKFTILKG 548  
Db 455 -----VXXMAAIKPIITYKG 469  
QY 549 IAAPLEKANVDTDAIIPKQFLKTIKRTGLGNALFYEMRFNEDGTEKSDFVLNKEPYRKAS 608  
Db 470 KIVPLFNDNIDTDQIIPKVHLKRIKSGFGPFADFWRYLPDGSNDPDPFNPKNKPYKAS 529  
QY 609 ILVCTGANFGCGSSREHAPWALNDFGIRSVIAPSFADIPFNNSFKNGLMPLPIPKDQAIE 668  
Db 530 ILI-TGDNFGCGSSREHAAWALKDYGFHIIAGSFSDIFYMNCTKNAMLPIVLEKNAR-E 587  
QY 669 AIAAEARAGKEIEVDLPNQLIKNATGETICTPEVEFRKHCLVNGLDDIGLTMQMEDKIA 728  
Db 588 HIAKYV-----EIEVDLPNQTV--SSPDKGFHFIDEITWKNKLVNGLDDIAITIOYESLIE 641  
QY 729 EFE 731  
Db 642 KYE 644

Search completed: March 17, 2003, 08:53:54  
Job time : 169 secs